

1/6

pXB296 around pos. 15,783

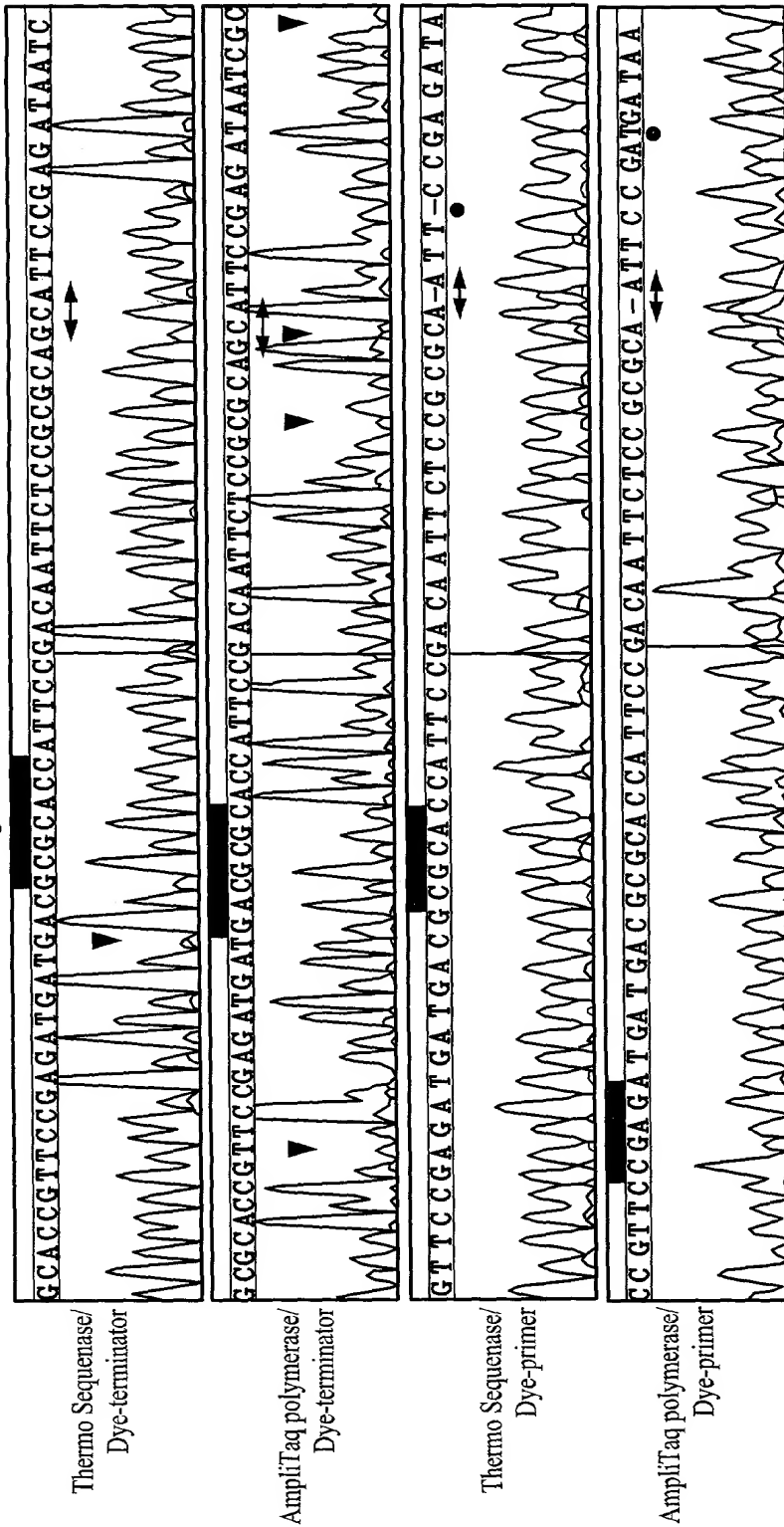
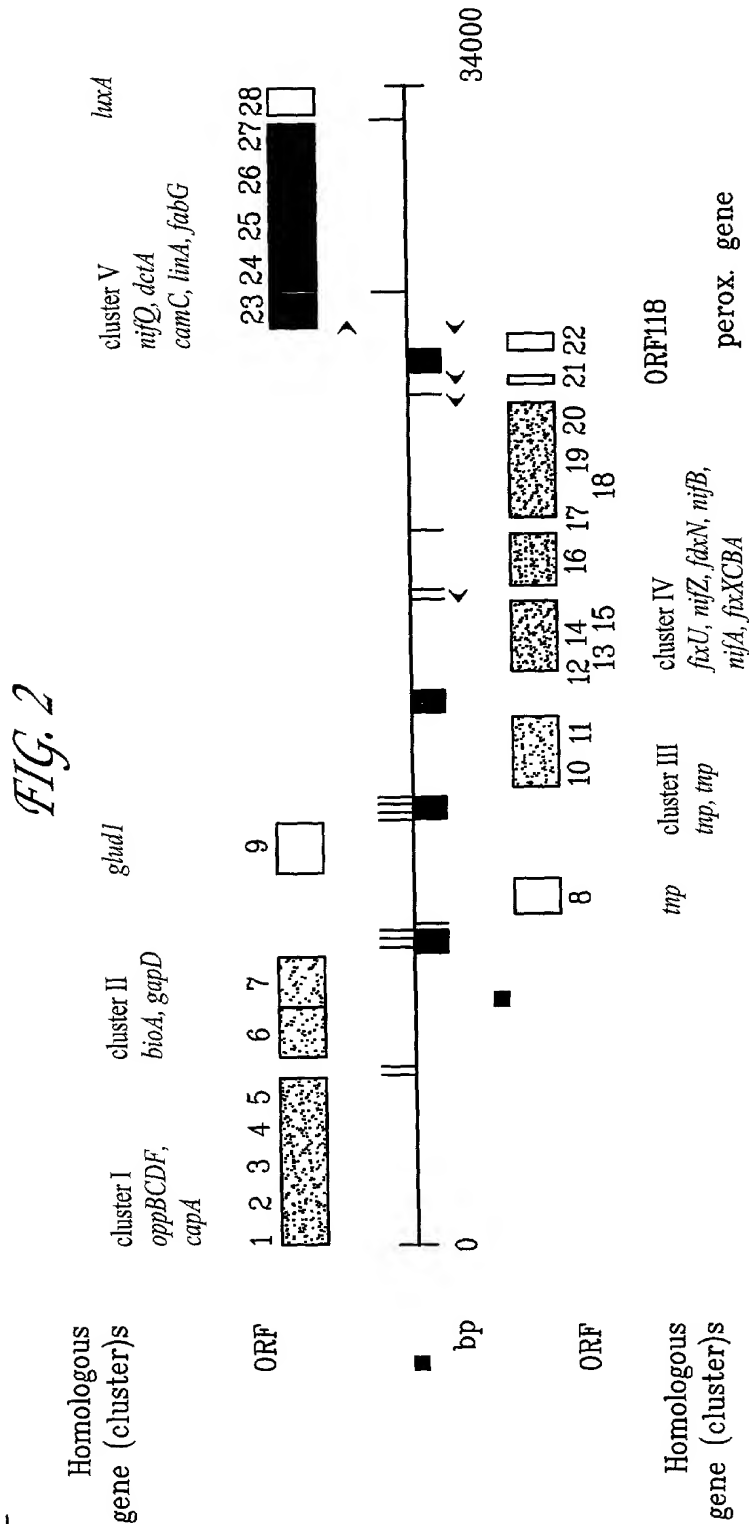


FIG. 1

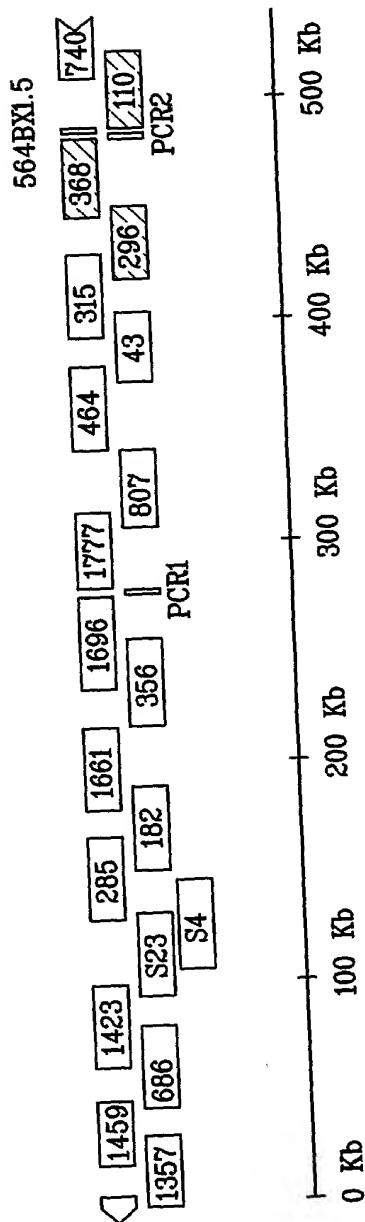
Comparison of sequences from pXB296 created by different cycle sequencing methods. The graphic outputs (program XGAP) of four electropherograms (traces) with the corresponding sequences generated by automatic base calling are shown around the pXB296 sequence position 15,783 (vertical line in the middle of each sequence). The readings cover a part of the minus strand of pXB296. (▼) Extremely low signals produced with dye terminators; (↔) the sequence GCA, which is compressed in dye primer scans; (●) automatic base-calling inaccuracies.

2/6



Organization of the predicted ORFs in pXB296 from *Rhizobium* sp. NGR234. Significant stem-loops (1)/stem-loop clusters (■), which might function as  $\rho$ -independent transcription terminators, are represented. Sequence motifs (open arrowhead), similar to  $\sigma^{54}$ -dependent promoter consensus sequences (TGGCACG-N<sub>4</sub>-TTTCG) + *nifA* upstream activator sequences (TGT-N<sub>10</sub>-ACA), are located at the following positions on the cosmid: 19,107 - 19,120 + 19,195-19,210; 24,787 - 24,800 + 24,878 - 24,893; 25,508 - 25,521 + 25,567 - 25,582; 26,815 - 26,828 + 26,941 - 26,956(all minus strand); 27,074 - 27,131 (several possibilities) + 26,969 - 26,984(plus strand).

FIG. 3



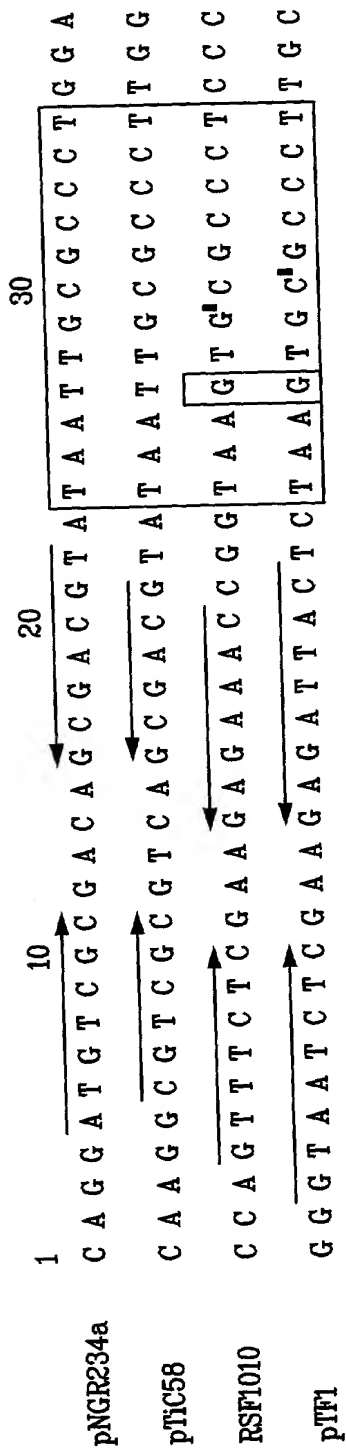
4/6

FIG. 4

	10	20	30
pRiA4b	C G C A A A A G A A A A G A G C C C C C T C A A C G T C G C C		
pNGR234a	C G C A A A A G A A A A A G G C C C C C A G A C G G T A A C C C		
pTiB6S3	C G C A A A A G A A A A A G G C C C C C G A A A C G - - G C G		
pRL8JI	C G C A A A - G A A A A A G C C C T C C G A A A C G G T G - G		
	40	50	60
pRiA4b	G T C G C G G A A G C C C T T C T G T C - T C - T C T A G C -		
pNGR234a	A T C G T G G A A G C C - - T C T C T C A T C G T T T A G C A		
pTiB6S3	T T C - C G G A A G A C C T T C T C T - A T A G T C T C G C A		
pRL8JI	T T C - C A G A A G G C C - T C T C T C A - - G T T T G G T C		
	70	80	90
pRiA4b	G C G A A C A G A A T C G C A T T T C C T C G A A T C C T C G		
pNGR234a	G C C T G - A G A A T C G C A T T T C C A C G A A T C G C A G		
pTiB6S3	G C T A A G A G A A T C G C A T T T T C A G G A A T C C C A G		
pRL8JI	G C T T A G A G A A T C G C A T T T C C C G G A A T C A C A G		
	100	110	120
pRiA4b	T C A A G A G T T T - - T T A G C G C C G T T T T G G - - T -		
pNGR234a	T C A A G A G T C T - - T T G G C A C C G G A A C G G - G T -		
pTiB6S3	T C A A G A G T C C C G T G A G G A A A G T A T C G T - T T C		
pRL8JI	T C A A G A G T C - - - - A A C G C C A C A C C G G C G T -		
	130	140	150
pRiA4b	- - - G A G C T G A T T T C C T T T T G C C T G C T G A A A G G		
pNGR234a	- - - G A G C G G A C G T C T T T T G C C T G A G G A T A G G		
pTiB6S3	G A C G A G C T G A T T T C T T T T G C C T A A C - A A A G G		
pRL8JI	- - - A G C C T - T T T C - T T T G C C T T G C G A A A G G		
	160		
pRiA4b	T G A A A G A T G -		
pNGR234a	T G A A A G A A G -		
pTiB6S3	T A C A A G G A - -		
pRL8JI	T G - A A G G A C A		

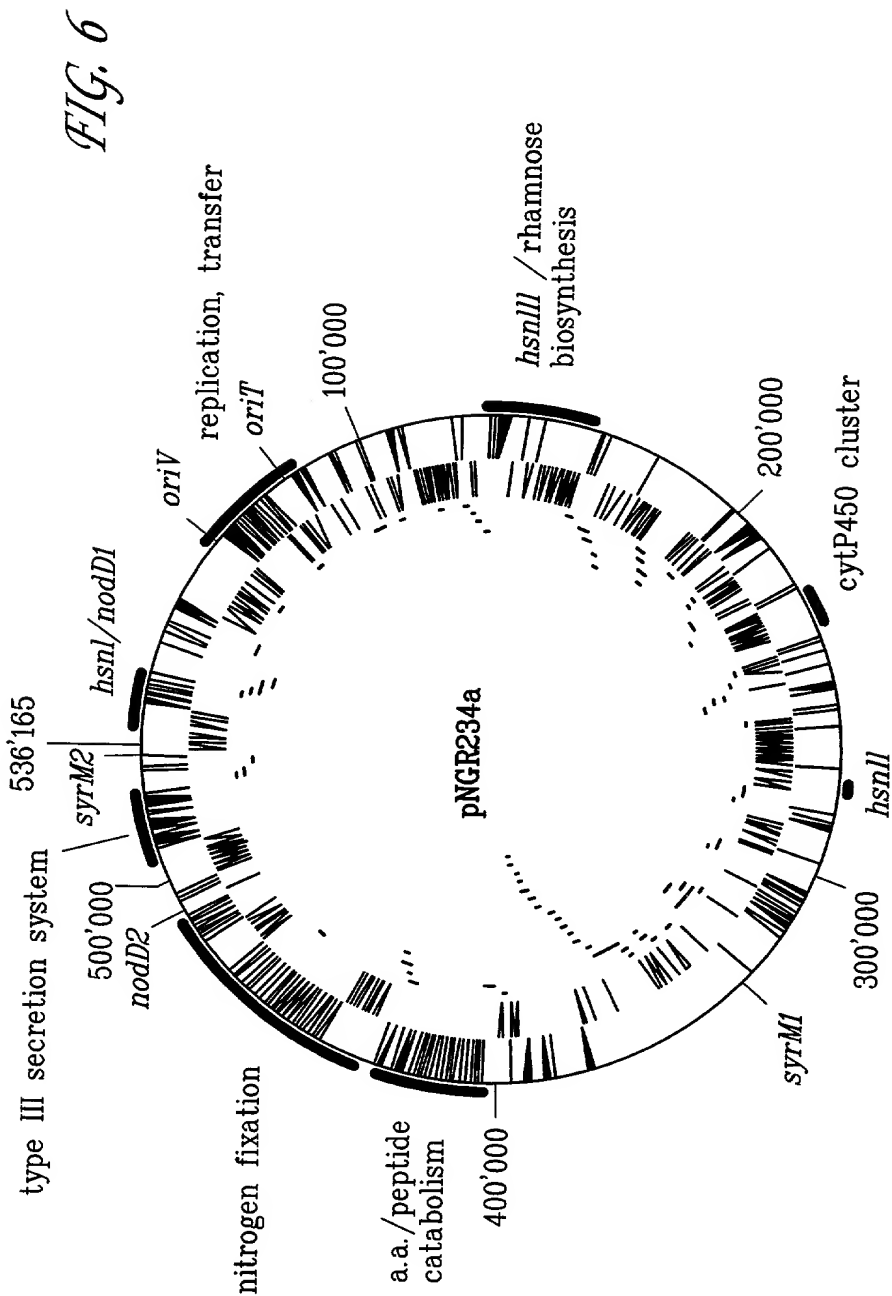
Multiple alignments of the nucleotide sequence of the replication origins of: the Ri plasmid of *Agrobacterium rhizogenes* (pRiA4b), the symbiotic replicon of NGR234 (pNGR234a), the Ti plasmid of *A. tumefaciens* BS63 (pTiB6S3) and pRL8JI of *R. leguminosarum* bv. *leguminosarum* (pRL8JI). Gaps introduced to give the best sequence alignments are marked by hyphens.

FIG. 5



Multiple DNA sequence alignments of the regions containing the origin of transfer of: the symbiotic plasmid of NGR234; pTiC58, the Ti plasmid of *A. tumefaciens* C58; RSF1010, a mobilisable plasmid of *E. coli*, and; pTFI, a mobilizable plasmid of *Thiobacillus ferrooxidans*. Major conserved nucleotide residues are boxed. Known "nick" sites corresponding to the nucleotide positions where the specific plasmid strand is cleaved are marked (z). Sequence features in the trailing portion of the *oriT* sites include inverted repeats which are marked by horizontal arrows.

6/6



Circular representation of the symbiotic plasmid of NGR234. Outer and inner concentric circles: coding regions identified on the plus and minus strands respectively. ORFs which belong to insertion-like elements are not visible. Thin concentric black lines represent mosaic sequences as well as complete or partial IS like repeats. Major gene clusters are highlighted as thick black concentric lines.